Topics in Advanced Clinical Trials Randomization, Blinding and Outcomes

Garnet Anderson, PhD
WHI Clinical Coordinating Center
Fred Hutchinson Cancer Research Center

Overview

- Randomization
  - Methods
  - Considerations
  - Subsampling
- Blinding
- Outcomes
  - Classifications in prevention trials
  - Data collection procedures
  - Data analysis issues
- Examples from WHI Clinical Trials

Randomization

- An unbiased method for assigning interventions to subjects
- Purpose: Assure intervention groups are comparable
- Achieved by assuring assignment of next subject is ‘unpredictable’
- Methods
  - Simple
  - Permuted block
  - Dynamic balancing

Simple randomization

- Probability of a subject being assigned to any particular group is independent of all other assignments
- Probability distributions are easily characterized
- Simple to implement
- Does not assure equally-sized groups
- Rarely used

Permuted blocks

- Random assignments for a group of subjects are generated jointly in predefined ratios
- Example: Block size 10 with 1:1 randomization
  1. Generate R[10,1]= vector of 5 ones and 5 zeros
  2. Generate U[10,1]= vector of 10 uniform (0,1) random numbers
  3. Create matrix M=R~U
  4. Sort M by column 2
  5. Assign group membership according to column 1

Permuted blocks

- Simple to implement
- Assures balance in treatment assignments within blocks
- Under staggered entry, provides balance over time
- Small block sizes may produce predictable assignments
- Large block sizes may lead to incomplete blocks and hence imbalance

Stratified permuted blocks

- Strata are defined by selected covariate information
- Permuted blocks are applied within these strata
- Assures balance in each stratum
- Increases the complexity of implementation
- Decreases the predictability of assignments
- Increases the chance of imbalance in overall numbers associated with incomplete blocks

Randomized permuted blocks

- Uses random-sized blocks to reduce predictability of later randomizations
- For example:
  - Select range of block sizes: 8,10,12,14,16
  - Generate random sequence of block sizes: 12,8,10,8,16 . . .
  - For each block in turn, generate a permuted block of randomization assignments

Randomized permuted blocks

- Straightforward to implement
- Virtually eliminates predictable assignments
- Chance of imbalance is a function of final block size

Dynamic balancing

- Designed for settings with several prognostic factors
- Uses measure of imbalance in selected covariates to determine probability of assignment to intervention group
- Focuses on balancing 'main effects'
  - Does not necessarily assure balance within cells defined by cross-classification
- Measure of imbalance can be tailored to emphasize specific covariates or subgroups

Pocock and Simon, Biometrics, 1975.

Dynamic balancing

- For each new subject, the characteristics are noted and the measure of imbalance between Arms A and B is calculated.
- If all relevant factors are currently balanced, Arm A is assigned with probability p=0.5.
- If assignment of next subject to Arm A would reduce imbalance, then randomize to A with p=p' where p' > 0.5, and Arm B with probability 1-p'.

Dynamic balancing

- Example: Assume there are 2 covariates:
- Age (< 50, 50+)
- Sex

with current allocation as shown. Let p' = 2/3.Define a measure of imbalance to be the sum of factor specific differences: ($\Sigma\Delta i$) = 8.

	A	B	Δ
F	12	13	1
M	15	11	4
<50	13	12	1
50+	14	12	2

If the next subject is M and <50, assign B with p =2/3.

If the next subject is F and <50, assign A with p = 1/2.

Selecting a randomization scheme

- Evaluation
  - Number and prevalence of prognostic factors
  - Strength of their association with outcome
  - Overall sample size and expected sample size within cells
  - Likelihood of investigators predicting subsequent randomization assignments
  - Logistics

Selecting a randomization

- Presence of clear prognostic factors suggests:
  - Stratification
  - Dynamic balancing, if expected sample size per cell is small
- Stratification/balancing on center is recommended for multicenter trials

Data analyses with structured randomization

- Linear models have a well-developed literature associated
- Generalizations to non-linear models are not direct
- In logistic or proportional hazards regression models, use of covariates in the model can be guided by their predictive strength

Subsampling

- Selecting a proportion of the overall trial cohort for specific tasks
- Used primarily for costly or burdensome data collection activities
- Examples:
  - Validation studies
  - Intermediate outcome studies
  - Secondary outcome studies requiring specific measurements

Subsampling

- Subsampling plan requires usual design considerations
- For prospective data collection, random selection can be done in conjunction with original randomization
- Blinding to membership in subsample may be needed
- May impact logistics, both positively and negatively

Randomization in the Women's Health Initiative

- Four randomized clinical trials
- Partial factorial design
- Participants may enroll in ≤ 3 trials, each requiring a separate randomization

Randomization in the Women's Health Initiative

- Separate randomizations for each trial
- Stratified, permuted block
- Stratification on
  - Clinical center site (49 sites)
  - Age (50-54, 55-59,60-69,70-79)
- Subsamples identified at baseline for
  - Ongoing blood collection and prospective analyses
  - 4 Day Food Records
  - Bone densitometry

Blinding

- The condition in which the randomization assignment is not revealed
- Purpose:
  - Preserve comparability of arms on all factors other than the intervention and its direct effects
  - In particular, assure unbiased outcomes ascertainment and adjudication

Blinded versus masked

Schulz and Grimes. Generation of allocation sequences in randomized trials: chance, not choice. Lancet 2002;359:515-519.

Blinding

- Variations
  - Single blinding: the participant is not informed of the randomization assignment
  - Double blinding: neither the participant nor the study staff interacting with participants are informed
  - Triple blinding: Double blinding with trial monitoring based on coded intervention arms.

Double blind versus single blind

Schulz and Grimes. Lancet 2002;359:515-519.

Blinding

- Reduces potential biases in all participant interactions and data collection, especially outcome ascertainment
- Feasibility depends strongly on the type of intervention
  - Most commonly implemented in drug studies
  - Only as effective as the placebo is comparable to the intervention on all aspects other than effect on disease
- Increases logistical complexity

Unblinding

- Revealing the randomization assignment
- Should be documented
- May be implemented in varying degrees
  - Clinical staff
  - Participant
- Preserve blinding of outcomes data collection process, whenever possible

Blinding and unblinding in WHI

- Computerized, blinded drug dispensing
  - Study database links participant to a unique bottle ID, based on randomization assignment
  - When bottle is retrieved, barcoded bottle ID is scanned into database to verify accuracy
- Official *unblinding* required for symptom management
- Supported by a database function
  - limited to authorized staff
  - self-documenting
- Unofficial unblinding from symptoms

Outcomes

- Most important data collection activity of a trial other than safety
- Deserving of considerable effort to assure data timeliness and quality
- Subject to considerable pressures from
  - Changing diagnostic methods
  - Changing medical-legal climate

Outcomes in prevention trials

- Usually diverse
- Observed only indirectly
- Require targeted efforts to ascertain, document and code
- WHI as an example

- Curb, McTiernan, Heckbert, Kooperberg, Stanford, Nevitt, et al. Outcomes ascertainment and adjudication methods in the Women's Health Initiative. Ann Epidemiol 2003:13 In press.

Outcomes/Endpoints

- Primary outcomes
  - Foundation of the trial
  - Drives the statistical design
  - Limited to a small number
- Secondary
  - Have noteworthy scientific interest
  - May have less preliminary data
  - Trial may not have sufficient power to answer definitively

Outcomes

- Safety outcomes
  - Known or suspected adverse effects
  - May need to be considered in trial design
  - Have a prominent role in trial monitoring

Outcomes

- Intermediate outcomes
  - A measurable quantity predictive of a clinical outcomes
  - Useful as
    - Proof of principle
    - Comparing effects in subgroups where there may be limited power for comparing clinical outcomes

Outcomes

- Surrogate
  - A measure or event that captures the full effect of the intervention on the disease outcome
    - Advantageous when ascertained more easily or earlier in the disease process
    - Rigorous statistical criteria for establishing surrogacy:

E{ Disease | Intervention, Surrogate }
= E{ Disease | Surrogate}

WHI primary & secondary outcomes

	DM	HRT	CaD
CHD	2°	1°	X
Angina	2°	2°	X
Revascularization	2°	2°	X
CHF	2°	2°	X
Peripheral vascular disease	2°	2°	X
Stroke	2°	2°	X
Venous thromboembolic disease	X	2°	X
Total CVD	2°	2°	X
Breast cancer	1°	1°S	2°
Colorectal cancer	1°	X	2°
Endometrial cancer	2°	2°	X
Ovarian Cancer	2°	2°	X
Total Cancer	2°	2°	2°
Hip Fractures	X	2°	1°
Other Fractures	X	2°	2°
Diabetes	2°	X	X
Total Mortality	2°	2°	2°

Outcomes ascertainment

- 1st priority: equal ascertainment across intervention arms
  - NOTE: Outcomes data collection can be blinded to randomization assignment, even in an otherwise unblinded trial.
- 2nd priority: complete ascertainment

Outcomes coding/adjudication

- Standardization always preferable for
  - Definitions
  - Documentation
  - Adjudication procedures
  - Adjudicators

WHI outcomes ascertainment

- Self-report of new clinical events collected at regular, protocol defined intervals (6 months)
  - Avoided non-routine reports for outcomes ascertainment to reduce potential for bias
    - Women with symptoms
    - Women in DM intervention arm
  - Self-report of safety outcomes could trigger processes to stop intervention

WHI outcomes ascertainment

- Search of National Death Index
  - Obtain date and cause of death information based on death certificates only
  - Substantial delay between date of death and appearance in the NDI
  - May not follow-up with additional requests to family or providers for documents
  - Value depends on adequacy of follow-up procedures and quality of personal identifiers
  - Consider providing names with known vital status (both deceased and alive) to estimate hit rates

Outcomes adjudication

- Classification of health events according to pre-defined criteria
- Criteria should include
  - Explicit definitions
  - Required documentation

WHI outcomes documentation

- Self-report of specified outcomes, or closely related ones, spawned a process of documentation and adjudication
  - Details of event were sought (e.g., dates and locations of hospitalizations)
  - Specific records required for each endpoint type
    - Path reports for cancers
    - ECGs and enzymes for MI

WHI outcomes adjudication

- Completed outcomes records provided to local clinic's physician adjudicator for review and coding.
- Central adjudication
  - All primary and safety outcomes
  - All deaths
  - Selected other endpoints (%)
    - Related to primary outcomes
    - Denied, self-reported outcomes

Outcomes data collection issues

- Timeliness of data collection
  - Critical for trial monitoring purposes
  - Important for adequate documentation
  - Difficult for bureaucratic reasons
    - Multiple institutions
    - Short interval medical release forms
    - Charges for records
    - HIPAA

Outcomes data collection issues

- Variation in documents received
  - Confusion in records requested
  - Differences in medical practice
    - Regional
    - Secular
  - Differences in aggressiveness of collection techniques

Outcome adjudication issues

- How many adjudicators per outcome?
- What defines agreement?
  - Primary diagnosis (e.g., invasive breast cancer)
  - Details of diagnosis (e.g., histology, grade, stage)
- What is the resolution process?

Outcomes data analysis issues

- Mapping outcomes to hypotheses
  - CHD is
    - Definite + probable MI
    - Coronary death
    - Silent MI

Outcomes data analysis issues

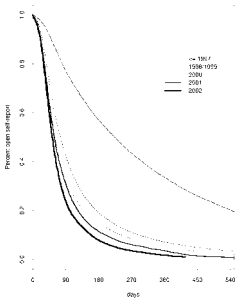
- Defining the “final” data
  - Local vs central adjudication
    - Central, if applied to all events
    - Local, if central not uniformly available
    - Unrefuted, all central + local that are not yet centrally adjudicated
  - Consideration of
    - Self-reports with no other documentation available
    - Passive data collection sources

Outcomes monitoring

- Timeliness, completeness, and accuracy of data collection
  - Self-report
  - Medical records retrieval
  - Local adjudication
  - Central adjudication

Timeliness of local adjudication

- Percent of self-reported events that have not yet been closed out through local adjudication by days since self-reported event data is received.



Performance monitoring for outcomes

- Performance Monitoring Committee
  - Regularly reviews clinic specific reports
  - Draws attention to performance issues
  - Offers assistance in systems design, tips for overcoming barriers
  - Membership drawn from Coordinating Center, NHLBI and well-performing clinics

Outcomes monitoring

- Rates of events (in control arm) relative to expected
  - Differences in recruited population
  - Healthy volunteer effect
  - Different
    - Outcomes ascertainment procedures
    - Diagnostic procedures
    - Outcomes definitions
  - Problems in the outcomes process

Summary

- Randomization
  - Several approaches available
  - May be tailored to assure objective of comparability is met
- Blinding
  - Helps preserves comparability
  - Should be implemented to the extent feasible within the design

Summary

- Outcomes
  - A critical data collections process
  - Requires planning, procedures, training, considerable effort, and ongoing monitoring